Center for Independent Experts (CIE) Independent Peer Review of benchmark stock assessments for black sea bass and witch flounder

SARC 62 Meeting, November 29 – December 2, 2016
Northeast Fisheries Science Center
Woods Hole, Massachusetts

Prepared for:
Center for Independent Experts

Prepared by:
Vivian Haist
1262 Marina Way,
Nanoose Bay, BC
Canada V9P 9C1

January 2017

1. EXECUTIVE SUMMARY

This report presents results of an independent peer review of the Northeast Regional Stock Assessment Review benchmark stock assessments of black sea bass and witch flounder (SARC 62), conducted for the Center for Independent Experts. The primary activity of the review was participation in the November 29 – December 2, 2016 SARC review in Woods Hole, Massachusetts.

The SARC 62 review process was thorough, effective, and resulted in a comprehensive review of the two stock assessments. The assessments had been developed by Stock Assessment Working Groups (SAW) who were thorough in developing and evaluating analytical assessment models and selecting a single model for the assessments, which facilitated the work of the Panel. The SARC 62 Panel reached consensus on all the assessment Terms of Reference.

SARC 62 reviewed a benchmark stock assessment for the black sea bass resource north of Cape Hatteras, NC. The previous black sea bass benchmark stock assessment had been rejected because of regional structure in the data which resulted in inadequate model fits. The current assessment is based on two spatial sub-units of the region, which alleviates many of the problematic model fits. The data sources for the assessment are adequate and used correctly. The catch history is relatively well documented, and regional and state bottom trawl surveys provide consistent time-series of recruit and adult abundance trends within each spatial sub-unit. Numerous sensitivity analyses were conducted to investigate alternative data and model assumptions, and the base model selected by the SAW Working Group for determining stock status was appropriate. Results from the two spatial sub-units were combined for determining stock status and for stock projections. Biological reference points, based on F_{MSY} and B_{MSY} proxies, were updated using results from the new assessment. The SAW Working Group conclusion that the stock is not overfished and overfishing is not occurring is consistent with the analyses presented. The black sea bass stock assessment provides a scientifically credible basis for developing fishery management advice.

Witch flounder occur primarily in the Gulf of Maine and are caught predominantly by commercial otter trawl. The data sources for the assessment appear to be adequate and used correctly. The analytical stock assessments, conducted using age-structured methods (ASAP and VPA), had major retrospective patterns, and the SAW Working Group proposed applying retrospective adjustments to fishing mortality and abundance estimates for stock status determination and stock projections. The retrospective pattern is a diagnostic for model misspecification, either in the data and/or structural assumptions of the model, and retrospective adjustment is an *ad hoc* procedure that may not result in appropriate parameter corrections. As such, the SARC 62 Panel rejected the witch flounder analytical stock assessment, a decision that I strongly support. The previously accepted VPA stock assessment, updated with the new data, also exhibited a major retrospective pattern, and was rejected by the SARC 62 Panel. As such, there is no analytical basis for updating the biological reference points or for determination of stock status. An alternative data-poor *empirical approach*, using swept-area biomass estimates from the NEFSC trawl surveys and results from a *sweep study* for calibration, could provide a scientifically defensible basis for developing fisheries management advice.

2. BACKGROUND

This document reports on an independent peer review of the benchmark stock assessments for black sea bass and witch flounder, conducted for the Center for Independent Experts. The primary activity of the review was participation at the November 29 – December 2, 2016 Northeast Regional Stock Assessment Review Committee (SARC) meeting in Woods Hole, Massachusetts.

The benchmark stock assessments were developed through the Northeast Stock Assessment Workshop (SAW) process, with separate SAW Working Groups comprised of regional experts for the black sea bass and witch flounder assessments. The objective of the SARC peer review is to determine whether the scientific assessments are adequate to serve as a basis for developing fishery management advice.

The 62nd SARC Panel was comprised of three CIE reviewers and a chair from the New England Fisheries Management Council Scientific and Statistical Committee (SSC) (Appendix 1).

3. DESCRIPTION OF REVIEW ACTIVITIES

The activities undertaken for this review are: 1) review and assimilation of background material and reports provided by the NMFS Project Contact prior to the SARC meeting, 2) participation in the SARC meeting, 3) contribution to the Panel Summary report, and 4) preparation of this report.

The materials provided to prepare for the SARC meeting included: draft stock assessment documents and assessment summaries for black sea bass and witch flounder; previous black sea bass and witch flounder assessment documents; and additional background documents summarizing research to support the assessments (Appendix 2).

The primary focus for the SARC Panel during the November 30 – December 2, 2016 meeting included:

- Determining whether data were adequate and used properly, the analyses and models were carried out correctly, and conclusions were reasonable and consistent with the analyses presented.
- Determining whether each stock assessment Term of Reference (ToR) was completed successfully.
- Determining whether the scientific assessments were adequate to serve as a basis for developing fishery management advice.
- Reviewing and agreeing text for the black sea bass and witch flounder stock assessment summary documents.

The SARC 62 *Statement of Work* (Appendix 3) defines the scope of this review.

This report, prepared for the CIE, reflects my own views which are consistent with the Panel's conclusions on all substantive issues, as described in the SARC 62 Summary report, summarizing the Panel's views and conclusions relative to the review ToRs, was prepared by Panel members during and after the meeting.

4. SUMMARY OF FINDINGS

4.1 OVERVIEW

The SARC 62 review process for black sea bass and witch flounder was thorough, effective, and resulted in a comprehensive review of the two stock assessments.

The SAW Working Groups had done a thorough job in developing and evaluating the assessment models, selecting a base case model for the assessment, and reporting the results. This greatly facilitated the work of the SARC 62 Panel.

The SARC 62 Panel made minor requests of the assessment scientists for additional background and supporting documentation, and some additional model runs to examine model behaviours. The Panel reached consensus on all assessment Terms of Reference and concluded that the black sea bass assessment provided a scientifically credible basis for developing fishery management advice. The Panel rejected the witch flounder analytical assessment model proposed by the SAW Working Group because it exhibited a major retrospective pattern. An alternative data-poor *empirical approach*, also put forward by the Working Group, should provide a scientifically credible basis for developing fishery management advice for witch flounder.

4.2 BLACK SEA BASS

SARC 62 reviewed a benchmark stock assessment for the black sea bass resource north of Cape Hatteras, NC. The previous black sea bass benchmark stock assessment had been reviewed and rejected by SARC 53 on the basis that there appeared to be regional structure in the data which resulted in inadequate fits to survey indices. The assessment reviewed by SARC 62 was based on two spatial sub-units of the region, which alleviated some of the problematic model fits.

The data sources for the assessment were adequate and in general used correctly. The catch history is relatively well documented, and regional and state bottom trawl surveys provide consistent time-series of recruit and adult abundance trends within each spatial sub-unit. A statistical catch-age model (ASAP) replaced the previously accepted length-based assessment model (SCALE). Numerous sensitivity analyses were conducted to investigate alternative data and model assumptions, and the base model selected by the SAW Working Group for determining stock status was appropriate.

The black sea bass assessment, based on separately modelling two spatial sub-units, had major retrospective patterns but in different directions for each sub-unit. Abundance trends for the two sub-units combined were very similar to those from models that assessed the stock as a single unit. Given the lack of retrospective pattern for the single area assessment and the similarity of abundance trends among single area and two area models, the assessment results from the *two area* model proposed by the SAW Working Group should be relatively robust to the uncertainty resulting from the retrospective pattern.

Biological reference points (BRPs), based on F_{MSY} and B_{MSY} proxies, were updated using results from the new assessment. The SAW Working Group conclusion that the stock is not overfished and overfishing is not occurring is consistent with the analyses presented.

Some suggestions about alternative approaches and assumptions for the stock assessment modelling that may improve future assessments are suggested below, but these are unlikely to change the assessment of stock status.

Findings relative to each of the Terms of Reference (ToR) follow:

1. *Summarize the conclusions of the February 2016 SSC peer review regarding the potential for spatial partitioning of the black sea bass stock. The consequences for the stock assessment will be addressed in TOR-6.)*

This ToR was completed.

The SSC peer review approved the SAW Working Group proposal to partition the northern black sea bass stock at approximately Hudson Canyon, creating northern and southern spatial sub-units for stock assessment purposes.

There is fairly strong support for this separation from tagging data, differences in the 2011 year-class signal, and different abundance trends in the northern and southern sub-units. The tagging data suggests that some black sea bass from the northern sub-unit are caught while over-wintering in the southern sub-unit, so the two sub-units are not completely independent.

While the spatial partitioning of the black sea bass stock should allow for improved fits of stock assessment models to the data, there is no basis for changing the current management of the resource as a single stock unit.

2. *Estimate catch from all sources including landings and discards. Characterize the uncertainty in these sources of data. Evaluate available information on discard mortality and, if appropriate, update mortality rates applied to discard components of the catch. Describe the spatial and temporal distribution of fishing effort.*

This ToR was completed satisfactorily.

The black sea bass commercial catch is primarily taken by otter trawl, pot, and hand-line gear. There is also a large recreational fishery using hook and line, which accounts for about half of the black sea bass catch.

For assessment modelling, the catch data are aggregated into trawl and non-trawl categories and separated into spring (Jan. – June) and autumn (July – Dec.) seasons. The time series fitted in the assessment model begins in 1989, so discard estimates are directly available from observer programs and don't need to be extrapolated. The assessment report provides estimates of the uncertainty in the commercial discards. No information on the likely accuracy of other components of the catch (commercial landings and recreational landings/discards) are provided, though it was suggested during the review, that given the high commercial value of black sea bass, the reported landings should be fairly accurate (i.e. no incorrect reporting of species).

The distribution of black sea bass has shifted northward, recently extending into the Gulf of Maine. The commercial and recreational catch has also shifted northward, with catches increasing in the northern sub-unit and decreasing in the southern sub-unit.

A comprehensive summary of the available information on discard mortality rates was provided. This included a number of experimental studies of black sea bass survival resulting from hook and line releases. Given the literature cited, the 15% mortality estimate assumed for hook and line caught black sea bass is appropriate. Also, the assumption of 100% mortality for trawl and sink gillnet commercial discards is warranted.

Catch-at-length estimates were developed using length frequency sampling by market category. Because of limited sampling, length frequency data were combined across the northern and southern sub-units, and occasionally required "borrowing" from other years when sample sizes within a market category were inadequate.

Catch-at-age for the commercial and recreational fisheries was generated by applying age-length keys (ALK – age-at-length data) to the catch-at-length. However, ALKs were only available for the commercial fishery in some years (primarily since 2008), so ALKs from fishery independent surveys (primarily NEFSC) were used for all the recreational catch-at-age and the commercial catch-at-age in some years. Using ALKs from other sources will lead to bias in the age compositions unless the age structure of the "borrowed" ALK is the same as the fishery to which it is applied. This will not be the case if there are the differences in the selectivity ogives or differences in the time of year that catch samples are taken.

3. *Present the survey data being used in the assessment (e.g., indices of abundance, recruitment, state surveys, age-length data, etc.). Investigate the utility of fishery dependent indices as a measure of relative abundance. Characterize the uncertainty and any bias in these sources of data.*

This ToR was completed satisfactorily.

In addition to the broad area Northeast Fisheries Science Center (NEFSC) trawl surveys, there are numerous inshore trawl surveys that capture black sea bass. The SAW Working Group decided to exclude all autumn surveys from their modelling, which is appropriate as black sea bass will have begun their offshore migration at that time and variable proportions of the stock would be available to the surveys.

The inshore surveys used in the assessment models include the Northeast Area Monitoring and Assessment Program (NEAMAP) trawl survey which spans most of the coast and numerous state trawl surveys that provide only local abundance estimates (VA, MD, DE, NJ, NY, CT, RI, and MA). For some of the state surveys, relative abundance indices were developed using statistical standardization approaches (GLM). Though few details were provided, it appears that the analytical approaches were correct. Further consideration might be given to which state indices warrant inclusion in the stock assessment models. Based on model residuals some appear to contain very little information about abundance trends.

A recreational catch rate index (REC CPA) was developed from the Marine Recreational Fisheries Statistics Survey data (MRFSS). The selection of fishing effort (trips) with the potential of catching black sea bass was based on species guild associations. This approach requires specifying a threshold for similarity coefficients to determine which species to include in the guild, and sensitivity to the threshold criterion would be useful to include in the analysis. The REC CPA likely provides a reasonably consistent index of abundance, as the areas fished have been stable over time.

The NEFSC surveys (spring and winter) are fitted as separate time series for the Albatross and Bigelow vessels. While this decreases the information content of the surveys, there are many indices that span the Albatross-Bigelow split, which will provide a basis for scaling the relative catchability of the two vessels.

Age structures are not collected for the state surveys, so ALKs from the NEFSC surveys were used to estimate the survey age compositions. Given the underlying age compositions will differ among the surveys and they occur at different times of year, this will likely lead to bias in age compositions. Some of the state surveys catch predominantly age 1 black sea bass (VA, MD, DE, NY), and these were distinguished using length cutoffs, which is appropriate for the first age class.

4. *Consider the consequences of environmental factors on the estimates of abundance or relative indices derived from surveys.*

This ToR was completed satisfactorily.

General additive models (GAM) fitted to catch rate data from the NEFSC spring trawl survey indicated significant relationships with ocean conditions. Juvenile catch rates (age 1) were related to temperature, salinity and shelf water volume suggesting oceanic conditions may be important in determining year-class strength. The predictive power of the relationship was weak, so it would be premature to include the relationship in assessment models.

5. *Investigate implications of hermaphroditic life history on stock assessment model. If possible, incorporate parameters to account for hermaphroditism.*

This ToR was completed satisfactorily.

One of the concerns for a protogynous hermaphrodite species is that imbalance in the sex ratio resulting from higher fishing mortality rates on males will cause a reduction in fertilization rates. A simulation study investigated the effects of reduced fertilization on alternative measures of population reproductive potential. The simulation results suggest that the most appropriate biological reference point (BRP) for spawning potential depends on the degree to which fertilization is reduced as the sex ratio is distorted. When the potential for reduced fertilization is weak or unknown, the combined male and female spawning biomass is the most appropriate BRP to capture spawning potential.

6. *Estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock), using measures that are appropriate to the assessment model, for the time series (integrating results from TORs-1,-4, & -5 as appropriate), and estimate their uncertainty. Include a historical retrospective analysis and past projection performance evaluation to allow a comparison with most recent assessment results.*

This ToR was completed satisfactorily.

The stock assessment was conducted using the Age Structured Assessment Program (ASAP), a statistical catch-age model which is available in the NOAA fisheries toolbox. ASAP is not structured for spatial (multi-area) stock assessments, so separate analyses were conducted for the northern and southern sub-units (*two-area model*). Sensitivity analyses were conducted assuming a single area (*overall model*) and for an *area-exchange* model. For the *area-exchange* model some of the winter-spring catch and survey abundance (NEFSC survey) was moved from the southern sub-unit to the northern sub-unit to account for the seasonal movement of black sea bass from the northern sub-unit to the southern sub-unit. For the two-area approaches, biomass metrics were combined and fishing mortality averaged to represent the entire stock.

Application of statistical catch age models to fisheries data requires a number of assumptions, primarily associated with data weighting and the form of fishery and survey selectivity ogives. The modelling assumptions adopted by the SAW Working Group were appropriate and consistent with international best practice. Data weighting to balance model residuals was based on the McAllister and Ianelli (1997) approach. For future assessments, it would be useful to conduct a sensitivity using the Francis (2011) weighting approach, as it tends to result in lower weights for the age-composition data.

The primary issue with the black sea bass assessment was a major retrospective pattern for the *two-area model*, the model proposed by the SAW Working Group as the basis for the assessment. The direction of the retrospective pattern differed between the two sub-units, with a pattern of biomass overestimation in the southern sub-unit and underestimation in the northern sub-unit. The retrospective pattern was not resolved for the *area-exchange model*; however, the *overall model* had only a very minor retrospective pattern.

The *overall* model might be preferred because of its improved retrospective pattern; however, there are other diagnostics where the *two-area model* had better performance. The two-area model was better able to fit most of the survey indices, in particular the increasing abundance trend in the northern sub-unit seen in the NEAMAP survey, the REC CPA index, and the state surveys (MA, RI, CT, and NY). The 2011 year-class is consistently very strong in northern area surveys and only average or lower strength in southern area surveys. Also, the relative catchability of the NEFSC Albatross survey to the Bigelow survey is much closer to its expected values (from the calibration study) for the *two-area model* than for the *overall model*.

Uncertainty in fishing mortality and abundance measures were estimated using the ASAP Markov-chain Monte-Carlo (MCMC) algorithm. This approach provides an estimate of uncertainty conditional on the structure of the assessment model. A series of ASAP model sensitivities were also run, and these provide an indication of the additional uncertainty resulting from model specification. The sensitivities included: the *overall model*, the *area-exchange model*, and a series of two-area models where each survey index was removed in turn. Results indicated relatively low sensitivity to the alternative model formulations.

An alternative age-structure model, Stock Synthesis (SS3), was also fitted to the black sea bass data. This model, though having many advantages over ASAP for the black sea bass assessment, was not considered as a potential base case by the SAW Working Group because there was insufficient time for its full exploration. The SS3 implementation was more consistent with the black sea bass data and perceived stock dynamics in that it directly models: multiple areas with movement; sex-specific dynamics; and fits directly to length frequency data and conditional age-at-length, as available. Results from the SS3 application were remarkably similar to those from the ASAP models, given there are very different assumptions and approaches between them.

Biomass and recruitment estimates for the black sea bass stock assessment are calculated as the sum of the retrospective adjusted estimates for the northern and southern sub-units, and fishing mortality as the average of the retrospective adjusted northern and southern sub-unit estimates. The retrospective adjustment is an attempt to account for the retrospective pattern, though it is an ad-hoc process and as such the degree to which it corrects estimates is unknown. However, given the retrospective patterns for the two sub-units are in opposite directions and to some degree cancel out, and the similarity in abundance trends for the combined northern and southern sub-units among the various model analyses, results of the stock assessment should be fairly robust to the retrospective uncertainty.

7. *Estimate biological reference points (BRPs; point estimates or proxies for B_{MSY}, $B_{THRESHOLD}$, F_{MSY}, and MSY), including defining BRPs for spatially explicit areas if appropriate, and provide estimates of their uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for BRPs. Comment on the appropriateness of existing BRPs and the "new" (i.e., updated, redefined, or alternative) BRPs.*

This ToR was completed satisfactorily.

Although the stock assessment is based on separate models for the northern and southern sub-units, the SAW Working Group appropriately argued that there is no evidence that the two sub-units represent distinct stocks and developed BRPs for the entire region.

The SAW Working Group selected a fishing mortality reference point of $F_{40\%}$ as a proxy for F_{MSY} as there was no evidence of a stock-recruitment relationship on which to base a direct estimate of F_{MSY}. The $F_{40\%}$ proxy is appropriate, given the published literature for Perciformes. $F_{40\%}$ was estimated as 0.355 and 0.365 for the northern and southern sub-units, respectively, for an average $F_{40\%}$ of 0.36 for the entire stock.

Biomass reference points were calculated from long-term simulations fishing at the F_{MSY} proxy, assuming the recent (2013 – 2015) fishery selectivity pattern and the 2000 – 2015 empirical recruitment estimates. The simulations were conducted for combined parameter estimates for the northern and the southern models for each of the MCMC iterations. The estimated uncertainty in the biomass reference points (SSB_{MSY}, B_{MS}Y, and MSY) reflects uncertainty in future recruitment and uncertainty in the recent fishing selectivity ogive. Additional simulations based on the *overall model* resulted in slightly higher biomass reference points, reflecting the additional uncertainty that arises when alternative model formulations are considered.

The approach used to estimate BRPs for black sea bass is consistent with what is used for other stocks in the region. The black sea bass application is slightly more complicated because of the need to combine information from two sub-units, but this does not deter from the applicability of the results.

8. *Evaluate overall stock status with respect to a new model or new models that considered spatial units developed for this peer review.*

This ToR was completed satisfactorily.

The retrospective adjusted estimates of 2015 *F*, *SSB*, and *B* for the northern and southern sub-units were combined (summed for biomass and averaged for fishing mortality) for comparison to BRPs. Given the major retrospective pattern for both sub-units, the retrospective adjustment is consistent with the approach adopted for other groundfish stocks in this region. The retrospective adjustment did not change conclusions about stock status.

The stock status for black sea bass is that the stock is not overfished and nor is it experiencing overfishing.

9. *Develop approaches and apply them to conduct stock projections.*
    a. *Provide numerical annual projections (3-5 years) and the statistical distribution (e.g., probability density function) of the OFL (overfishing level) that fully incorporates observation, process and model uncertainty (see Appendix to the SAW TORs). Each*

projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. Use a sensitivity analysis approach in which a range of assumptions about the most important uncertainties in the assessment are considered (e.g., terminal year abundance, variability in recruitment, and definition of BRPs for black sea bass).

b. *Comment on which projections seem most realistic. Consider major uncertainties in the assessment as well as the sensitivity of the projections to various assumptions.*

c. *Describe this stock's vulnerability (see "Appendix to the SAW TORs") to becoming overfished, and how this could affect the choice of ABC.*

This ToR was completed satisfactorily.

Short term stock projections were conducted for the northern and southern spatial sub-units using the AGEPRO software (available in NOAA toolbox) and results combined to reflect the entire stock.

Projections assumed: fishing at the F_{MSY} proxy (or alternatively, $F_{StatusQuo}$); recruitment sampled from the empirical 2000 – 2015 distribution; recent (2013 – 2015) average weight-at-age and selectivity ogives; and 2016 catch equal to the ABC (or 20% higher) and the recent north-south catch split. Uncertainty was incorporated in the stock projections by sampling from terminal year MCMC estimates of number-at-age. Projections were conducted with and without adjustment for the retrospective pattern.

A sensitivity stock projection was conducted for the *overall model*, which showed similar results to the retrospective adjusted combined area models. As such, the SAW Working Group suggested that the retrospective adjusted area combined results provide reasonable projection estimates.

Projection estimates were provided in terms of the means, and the means plus and minus two standard deviations. Results could be presented in terms of probabilities of exceeding thresholds BRPs for fishing mortality or falling below threshold BRPs for biomass, as requested in the ToR.

The stock assessment, and hence stock projections, did not explicitly account for hermaphroditism beyond using male and female spawning stock biomass. The similarity of the stock assessment biomass trend with that from the SS3 model, which did model sexes separately, suggests that hermaphroditism is unlikely to have a large impact on assessment results.

10. *Review, evaluate and report on the status of the SARC and Working Group research recommendations listed in recent SARC reviewed assessments and review panel reports. Identify new research recommendations.*

This ToR was completed satisfactorily.

The SAW assessment document lists research recommendations from the SARC 53 benchmark assessment and reports on relevant results. SARC 53 rejected the proposed stock assessment and one of the key recommendations was to develop a spatially structured assessment model that allows mixing. Significant progress was made for the SARC 62 black sea bass assessment to account for spatial structure in this stock.

The SAW Working Group provided a number of research recommendations for black sea bass including: expanded genetic studies with smaller spatial increments; investigate the impact of climate change on black sea bass; evaluate sex change and sex ratio; investigate black sea bass catchability in a variety of survey gear; investigate social and spawning dynamics; additional work to study habitat use and seasonal changes; and investigate using samples collected by industry study fleets.

In addition to the recommendations from the SAW Working Group, I suggest a few additional research areas that could substantially improve future black sea bass stock assessments.

Further work should be conducted using the SS3 modelling platform, or one with similar capabilities, to reconstruct the black sea bass population. SS3 has the capacity to directly model many of the important features of black sea bass dynamics and the types of data that are available. Aspects that I think are important to model are: spatial structure with movement; sex-specific and explicit sex change; and fitting to length frequency data directly where no directed age data are available.

Compile and report on all available sex ratio data for the black sea bass stock with the objectives of determining if there is evidence for changes in sex ratio at age or length over time, and determining if any of the fisheries are sex selective.

The observed retrospective pattern in the *two area model* approach suggests some form of model misspecification, which warrants further research. One consideration, consistent with many of the observations, is that there has been a general northward movement in the stocks distribution.

4.3 WITCH FLOUNDER

SARC 62 reviewed a benchmark stock assessment of the US witch flounder resource. Witch flounder occur primarily in the Gulf of Maine and are caught predominantly by commercial otter trawl. Stock assessment analyses were conducted using age-structured methods, including the ASAP model and VPA. The age-structured approaches, including the proposed ASAP assessment base case, had major retrospective patterns whereby it appears that biomass is consistently overestimated and fishing mortality underestimated. The SAW Working Group had proposed applying retrospective adjustments to fishing mortality and abundance estimates for stock status determination and stock projections, as has been done for a number of Northeast groundfish stocks. The retrospective pattern is a diagnostic that clearly indicates some form of model misspecification, either in the data and/or structural assumptions of the model. The retrospective adjustment is an *ad hoc* procedure and it may not result in the appropriate corrections of model parameters. The SARC 62 Panel rejected the witch flounder analytical stock assessment, a decision that I strongly support. The previously accepted VPA stock assessment, updated with the new data, also exhibited a major retrospective pattern and was rejected by the SARC 62 Panel. As such, there is no analytical basis for updating biological reference points or for determination of stock status.

The SAW Working Group presented results from an *empirical approach*, which could provide a basis for determining overfishing limits for witch flounder. The *empirical approach* estimates swept-area biomass from the NEFSC trawl surveys, using results from the *sweep study* to calibrate to absolute abundance. Exploitation rates are calculated as the ratio of catch to biomass (adjusted for exploitable fraction). While there is considerable uncertainty in the swept-area biomass estimates, catch levels resulting from application of a target exploitation rate to survey biomass should be relatively robust to those uncertainties.

The data sources for the witch flounder assessment appear to be adequate and used correctly. The retrospective patterns in catch-age analyses raise the concern of significant catch underreporting that should be investigated. The long-term NEFSC trawl surveys and ASMFC shrimp survey provide solid information on witch flounder stock trends.

Findings relative to each of the Terms of Reference follow:

1. *Estimate catch from all sources including landings and discards. Describe the spatial and temporal distribution of landings, discards, and fishing effort. Characterize the uncertainty in these sources of data.*

This ToR was completed successfully.

Commercial landings of witch flounder declined sharply from a peak of over 6,000 mt in 1982, were relatively stable at 2,000 – 3,000 mt until 2003, and declined steadily after that to a historic low level of 662 mt in 2015. The commercial fishery for witch flounder is conducted primarily by otter trawl in the Gulf of Maine, with the majority of the catch landed in Maine and Massachusetts. Estimates of uncertainty in the commercial landings were not provided. The possibility of under-reporting of landings is posed as one of the potential factors in the stock assessment retrospective pattern (discussed below), and as such warrants future work to determine if under-reporting can be quantified.

Extensive work was conducted by the SAW working group to refine the methods to quantify commercial discards and to extend the estimates to additional gear sectors. The methodologies employed were appropriate, and the resulting estimates of witch flounder discards and the uncertainty in the estimates should be fairly reliable.

Length frequency and age-at-length sampling of commercial landings is adequate over the assessment time frame (1982- 2015). The methods used to estimate the age compositions of landings is appropriate and results in relatively precise estimates of landed numbers-at-age (c.v.s generally < 15%).

For some of the gear sectors, the length frequency samples of discards were inadequate to characterize the discards and length frequency samples from other sectors or years were used to fill the gaps. The criteria for doing this are appropriate and likely result in fairly accurate estimates of the discard length frequencies. Because age-at-length samples are not collected for the discard component of the catch, age-length keys from the NEFSC spring and autumn surveys were used to estimate the discards-at-age. This has the potential to result in biased age compositions for discards when the age-composition of discards is different than the survey or if length-at-age differs because of seasonal timing differences. These biases will likely have only a minor effect on the age-composition of the catch.

There has been a strong truncation in the age structure of the commercial catch since the late 1980s. A slight expansion, with greater numbers of older fish, occurred during the early 2000s, but this was followed by a further reduction in the numbers of older fish.

2. *Present available federal, state, and other survey data, indices of relative or absolute abundance, recruitment, etc. Characterize the uncertainty and any bias in these sources of data and compare survey coverage to locations of fishery catches. Select the surveys and indices for use in the assessment.*

This ToR was completed satisfactorily.

One of the primary abundance time series for witch flounder are the NEFSC spring and autumn trawl surveys. Abundance estimates for the Bigelow component of the time series are adjusted to Albatross-equivalent units using a calibration factor from the 2008 dual fishing study. The calibration uses a single factor, although there is some indication that length-based adjustments would be more appropriate. The survey strata used for estimating survey abundance indices appears to be appropriate in that they account for most of the surveyed witch flounder abundance and are consistent with the extent of the commercial fishery.

Results of a twin-trawl study (*sweep study*) to estimate the maximum catchability of the Bigelow survey trawl are an integral component of the witch flounder stock assessment. This study compared catch rates of the Bigelow rockhopper trawl gear with that of a chain sweep net designed to maximize flatfish catch. The experiment was well designed and the data analysis was thorough and comprehensive. The primary assumption required to estimate the maximum catchability (q) of the Bigelow gear was that the chain sweep gear was 100% efficient between the wings. This assumption implies that there was no herding effect of the chain sweep gear, which appears to be a reasonable assumption given the gear and survey design. However, when the maximum catchability estimates were assumed in catch-age models the fits to model data were extremely poor, indicating contradictions between the maximum q estimate and other model data and/or assumptions (see ToR 4).

Truncation in the age structure is seen in the NEFSC surveys, as in the commercial catch, though with a more pronounced effect in the survey data.

The Atlantic States Marine Fisheries Commission (ASMFC) summer shrimp survey in the Gulf of Maine is a relatively long time series (1984 onward) that consistently catches witch flounder. Although the spatial coverage of the survey is limited, it does encompass most of the range of juvenile witch flounder. Abundance indices from this survey are fitted in the catch-age models assuming ages 3 and 4 are fully selected. Witch flounder age-at-length data are not collected for this survey, so age-length keys from the NEFSC spring and fall (combined) surveys are used to estimate the age-structure. This will likely result in some bias in the age compositions because of different selectivity and differences in the timing of the two surveys (the variance of length-at-age will be higher for the combined NEFSC surveys than for the ASMFC survey). These biases may not be large, but a preferable approach is to fit the length frequency data directly in a model.

The Maine Department of Marine Resources and New Hampshire Fish and Game Department (MENH) spring and fall bottom trawl surveys has limited spatial coverage, but much of the juvenile witch flounder range is covered. As for the ASMFC survey, no ageing is conducted for this survey, so NEFSC age-length keys were used to estimate the age compositions. Similar caveats about potential bias from the application of age-length keys from surveys with different age-compositions apply.

A number of other surveys were considered for inclusion in the witch flounder stock assessment. The SAW Working Group appropriately decided that due to survey limitations, primarily limited spatial coverage, none of these were appropriate to use in the assessment models.

A fishery dependent landings-per-unit-effort (LPUE) index was developed for potential inclusion in the stock assessment models. Analyses focused on using the commercial Dealer landings data because it had an extended time series and large and synoptic samples. A sub-set of the data was selected based on a lower threshold for the proportion of witch flounder in the total landings for a trip. Thresholds examined were 10%, 25%, and 40%. LPUE indices were calculated using a standard GLM approach. While overall trends were similar among the indices calculated for the different thresholds, the lower thresholds (in particular 10%) resulted in a greater increase following the historic low levels in the 1990s. Compared with the NEFSC trawl survey indices, the LPUE indices are generally lower prior to the mid-1980s and higher since the early 2000s. That is, the LPUE time series has much less contrast than the fishery independent survey, and current indices are at least double the low levels of the 1990s. In contrast, the NEFSC survey indices are currently as low as the low levels in the 1990s. Appropriately, the LPUE index (40%) is only included in the assessment model for a sensitivity run. There are many reasons why time series of commercial fishery catch rates may not reflect stock abundance, including; changes in fishery regulations, changes in fisher behavior, and changes in fishing technology.

Catch curve analyses were conducted following cohorts in the NEFSC surveys and in the commercial catch. While this analysis is useful and informative when applied to the survey data, it is less so for the commercial fishery because results will be confounded by changes in effort. The survey catch curve analysis assumed ages 5 to 9 were fully selected. While Z estimates were highly variable, there was no indication of a time trend in the estimates which averaged about 0.4 to 0.5. Residuals from the linear catch curve fits were calculated to infer age-selectivity. This approach as applied to the survey data is appropriate, and did not provide evidence for domed survey selectivity.

3. *Investigate effects of environmental factors and climate change on recruitment, growth and natural mortality of witch flounder. If quantifiable relationships are identified, consider incorporating these into the stock assessment.*

This ToR was completed satisfactorily.

Analyses of the NEFSC trawl survey data indicate that while there have been changes in the distribution of juvenile and adult witch flounder, potentially related to changes in Gulf of Maine mean water temperatures, there is no evidence that movement has been out of the stock definition strata.

Analyses found no evidence for a relationship between temperature and recruitment.

Continued work to explore the effects of environmental factors on stock dynamics, in particular natural mortality, is warranted. This work would likely be more productive in a global context that considered other groundfish species occupying the same ecosystem.

4. *Estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series (integrating results from TOR-3 if appropriate), and estimate their uncertainty. Include a historical retrospective analysis to allow a comparison with previous assessment results and previous projections. Compare F's and SSB's that were projected during the previous assessment to their realized values.*

This ToR was completed satisfactorily. However, the analytical model selected by the SAW Working Group to represent the stock assessment was rejected by the SARC Panel, so reporting of some of these metrics is not possible.

A number of approaches for estimating fishing mortality and stock biomass were considered by the SAW Working Group who recommended results from an age-structured model (ASAP) as the basis for the stock assessment. The base model configuration had the following characteristics: fitting to the NEFSC swept-area estimates of biomass with the Albatross and Bigelow components treated as a single series, but estimating catchability; flat-top selectivity for both the commercial fishery and NEFSC surveys; fitting to the ASMFC shrimp and ME/NH trawl survey indices; three fishing selectivity time blocks; and reweighting of data sets to balance residuals.

The ASAP model configuration was appropriate given the data and ancillary information available and the application was correct. However, the model exhibited a major retrospective pattern whereby the stock biomass appears to be overestimated and fishing mortality underestimated. This retrospective pattern suggests some form of model misspecification, either errors in the data and/or model structure. The magnitude of the retrospective pattern is such that model results should not be considered to provide a credible scientific basis for management advice.

The ASAP base assessment had other issues that would in themselves not have been reason for rejecting the model, but corroborate that there is some form of model misspecification. These included: strong and common patterns in residuals (positive residuals for most of the final 15 years of the time series) for the fits to the NEFSC and ASMFC abundance survey indices and catchability (q) estimates of about 4 for the NEFSC surveys which had an expectation of 1 based on the *sweep study*.

Retrospective adjustment (rho adjustment) of terminal year abundance and fishing mortality estimates was proposed by the SAW Working Group as a method to correct for the retrospective bias in these parameters. However, this is an ad hoc procedure which cannot be relied on to provide appropriate management advice.

The retrospective pattern was not corrected in any of the ASAP model sensitivity trials. The sensitivity analyses evaluated one-off changes that included: domed-shaped selectivity in the fishery or the surveys; inclusion of the fishery-dependent LPUE time series; and down-weighting the age-composition data series fitted in the model (Francis weights). The assumptions of domed selectivity had minimal effect on model runs, likely because these were conducted separately for the fishery and survey (NEFSC) indices. While inclusion of the LPUE series and down-weighting age-composition data resulted in somewhat higher abundance estimates, these runs did not resolve the retrospective pattern, improve the problematic survey residual patterns, or provide q estimates close to their expectation of 1.

A series of exploratory runs were conducted to determine what alternative levels of natural mortality (M) or catch that would be required to eliminate the retrospective pattern. These indicated that 2.5 -3 fold increases in M or 3-5 fold increases in catch would eliminate the retrospective pattern and these were considered implausible.

The previously accepted age-structured VPA model was updated with data through 2015. This model also exhibited a major retrospective pattern, and so does not provide an acceptable alternative to the ASAP base case.

None of the analytical age-structured models presented provides an acceptable assessment as they either had major retrospective patterns or reflected unacceptable assumptions. Additionally, some of the alternative models had quite different abundance estimates or trends than the base model indicating the results were not robust to the uncertain assumptions of the analyses.

Results from an alternative data-poor method, the *empirical approach*, were also considered by the SAW Working Group. This approach calculates stock abundance directly from the NEFSC survey, incorporating results from the *sweep study* to estimate absolute biomass. Although there will be considerable (and unquantifiable) uncertainty in biomass estimates from this approach, the abundance trends should be considered relatively reliable. This approach indicates stock biomass declined since 2002, though it appears to have stabilized in recent years.

5. *State the existing stock status definitions for "overfished" and "overfishing". Then update or redefine biological reference points (BRPs; point estimates or proxies for B_{MSY}, $B_{THRESHOLD}$, F_{MSY}*

and MSY) and provide estimates of their uncertainty. If analytic model- based estimates are unavailable, consider recommending alternative measurable proxies for BRPs. Comment on the scientific adequacy of existing BRPs and the "new" (i.e., updated, redefined, or alternative) BRPs.

The SAW Working Group completed this ToR satisfactorily. However, the SARC rejected the proposed analytical assessment model and no acceptable alternative analytical results are available, so it is not possible to report on some of the BRPs.

The existing BRPs, estimated for the 2015 stock assessment update, were based on a VPA that was also rejected by the SARC 62 peer review. The VPA also had a major retrospective pattern, which makes it unreliable for status determination.

An alternative F_{MSY} proxy, based on the *empirical approach*, was suggested by the SAW Working Group as a back-up should the analytical assessment be considered indefensible. The *empirical approach* calculates beginning of year abundance as the average of the NEFSC autumn and spring trawl survey swept-area biomass, using the *sweep study* results to convert catch rates to absolute abundance. Exploitation rates are calculated as the ratio of catch to beginning year biomass. The SAW Working Group suggested that the average exploitation rate of the most recent 9 years (0.05) would be an appropriate proxy for F_{MSY} given that abundance has been relatively stable over that period. While there is considerable uncertainty in the absolute scale of abundance and exploitation rate estimates, the application of an exploitation rate target to survey-based abundance estimates will be relatively robust. While the use of an F_{MSY} proxy based on the empirical approach is *ad hoc*, the procedure is defensible.

6. *Evaluate stock status with respect to the existing model (from previous peer reviewed accepted assessment) and with respect to a new model (or possibly models, in accord with guidance in attached "Appendix to the SAW Assessment TORs") developed for this peer review. In both cases, evaluate whether the stock is rebuilt .*
    a. *When working with the existing model, update it with new data and evaluate stock status (overfished and overfishing) with respect to the updated BRP estimates.*
    b. *Then use the newly proposed model (or possibly models, in accord with guidance in "Appendix to the SAW Assessment TORs") and evaluate stock status with respect to "new" BRPs and their estimates (from TOR-5).*

The SAW Working Group completed this ToR satisfactorily. However, both the existing VPA model and the proposed ASAP model (and alternative analytical models) were rejected, so it is not possible to evaluate stock status.

The empirical area swept method does not provide a biomass threshold, but does indicate that the stock is at low historical levels.

7. *Develop approaches and apply them to conduct stock projections.*
    a. *Provide numerical annual projections (3 years) and the statistical distribution (e.g., probability density function) of the catch at F_{MSY} or an F_{MSY} proxy (i.e. the overfishing level, OFL) (see Appendix). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. Use a sensitivity analysis approach in which a range of*

assumptions about the most important uncertainties in the assessment are considered (e.g., terminal year abundance, magnitude and variability in recruitment).

b. *Comment on which projections seem most realistic. Consider the major uncertainties in the assessment as well as sensitivity of the projections to various assumptions. Identify reasonable projection parameters (recruitment, weight-at- age, retrospective adjustments, etc.) to use when setting specifications.*

c. *Describe this stock's vulnerability to becoming overfished, and how this could affect the choice of ABC. The choice takes scientific uncertainty into account (see Appendix).*

The SAW Working Group completed this ToR satisfactorily. However, the analytical assessment was rejected, so it is not possible to report on stock projections.

8. *Evaluate the validity of the current stock definition, taking into account what is known about migration, and make a recommendation about whether there is a need to modify the current stock definition for future stock assessments.*

This ToR was completed satisfactorily.

No new information was available for evaluating stock structure. Previously available information supports the current stock structure. Survey data suggests witch flounder on the western Scotian Shelf are discrete from those in the Gulf of Maine/Georges Bank, and deepwater witch flounder have slower growth and different otolith structure.

9. *Review, evaluate and report on the status of research recommendations from the last peer reviewed benchmark stock assessment. Identify new research recommendations.*

This ToR was completed satisfactorily.

The SAW assessment document lists research recommendations since the last benchmark assessment and reports on relevant results.

The SAW Working Group provided a number of research recommendations specific to witch flounder, including: aging archived structures from the MENH trawl survey; conducting a stock identification study; conducting larval growth studies; conducting tagging studies directed to estimating M; and further work on habitat preference to use in swept-area expansions. In addition, some of SAW Working Group research recommendations are generic and applicable to all SARC stock assessments. These include; additional options for fitting age compositions and assessment of their influence in catch-age models; determine veracity of reported catch statistics and potential for unreported catch; investigate how M might change over time; and develop geostatistical approaches for utilizing survey data.

I endorse the SAW Working Group recommendations, and in particular support work to investigate the retrospective pattern issue. There are major retrospective patterns for a number of Northeast groundfish stocks, suggesting there are likely common causes among them. A research project to investigate all stocks that exhibit this pattern is much more likely to be successful than a species-by-species approach that occurs only when a benchmark assessment is due.

5. SUMMARY AND CONCLUSIONS

The SARC 62 review process for black sea bass and witch flounder was thorough, effective, and resulted in a comprehensive review of the data and analytical methods used to assess the stocks. The working group process (SAW) used to develop the assessments ensured that assumptions and uncertainties in the assessments had been thoroughly investigated. Preparation of draft assessment documents and summary reports prior to the review meeting and limiting the review to two assessments ensured there was adequate time to become familiar with the assessment data and methods, and to thoroughly investigate implications of modelling assumptions. The terms of reference for the review process are clear and explicit, and provide a useful guide for the review.

The Review Panel members agreed on all substantive issues and the SARC 62 Summary Report represents consensus opinion. Stock assessment Summary Reports, agreed during the review, reflect the best possible use of the available information and conclude that the black sea bass resource is not overfished and overfishing is not occurring. The witch flounder analytical stock assessment was rejected, as was the previously accepted VPA model, so there no basis for developing biological reference points and reporting stock status. An alternative data-poor *empirical approach* could provide a scientifically credible basis for developing fisheries management advice.

6. REFERENCES

Francis, R.I.C.C. 2011. Data weighting in statistical fisheries stock assessment models. Can. J. Fish. Aquat. Sci. 68: 1124-1138.

McAllister, M.K. and J. N. Ianelli. 1997. Bayesian stock assessment using catch-age data and the sampling-importance resampling algorithm. Can. J. Fish. Aquat. Sci. 54: 284-300.

Appendix 1: SARC 62 Panel Members

Patrick Sullivan, New England Fishery Management Council SSC (Chair)
Anders Nielsen, CIE reviewer
Neil Klaer, CIE reviewer
Vivian Haist, CIE reviewer

Appendix 2: Bibliography of materials provided for review

Black Sea Bass

Background Papers

Blaylock J, Shepherd GR. 2016. Evaluating the vulnerability of an atypical protogynous hermaphrodite to fishery exploitation: results from a population model for black sea bass (*Centropristis striata*). Fish Bull. 114:476–489.

Brooks EN et al. 2008. Stock assessment of protogynous fish: evaluating measures of spawning biomass used to estimate biological reference points. Fish Bull. 106:12–23.

Keigwin B, Shepherd GR, Wuenschel MJ. 2016. Geomorphometric analysis indicates overlap in body shape between sexes of black sea bass (*Centropristis striata*). US Dept Commer, Northeast Fish Sci Cent Ref Doc. 16-07; 26p.

Miller AS, Shepherd GR, Fratantoni PS. 2016. Offshore Habitat Preference of Overwintering Juvenile and Adult Black Sea Bass, *Centropristis striata*, and the Relationship to Year-Class Success. PLoS ONE 11(1): 19p.

Moser J, Shepherd GR. 2009. Seasonal Distribution and Movement of Black Sea Bass (*Centropristis striata*) in the Northwest Atlantic as Determined from a Mark-Recapture Experiment. J Northw Atl Fish Sci. 40: 17–28.

Nieland JL, Shepherd GR. 2011. Comparing Black Sea Bass Catch and Presence Between Smooth and Structured Habitat in Northeast Fisheries Science Center Spring Bottom Trawl Surveys (Working Paper for SAW 53). 7p.

Northeast Fisheries Science Center. 2012. 53rd Northeast Regional Stock Assessment Workshop (53rd SAW) Assessment Summary Report. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 12-03; 33p.

Shepherd, G., K. Shertzer, J. Coakley, and M. Caldwell (Editors). 2013. Proceedings from a workshop on modeling protogynous hermaphrodite fishes. Raleigh, NC. 33p.

Working Papers

Fay G. 2016. Retrospective analysis for Black sea bass Stock Synthesis model 'run_164'. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 6p.

Fay, G and Cadrin S. 2016. Simulation testing assessment models for Black Sea Bass. Appendix A6, Stock Assessment Report of Black Sea Bass. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 26p.

Fay G, McNamee J, Cadrin S. 2016. Stock Synthesis Application to Black Sea Bass. Appendix A9, Stock Assessment Report of Black Sea Bass. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 59p.

Fay G, McNamee J, Cadrin S. 2016. Stock Synthesis Application to Black Sea Bass. Appendix A9, Stock Assessment Report of Black Sea Bass. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 59p.

Robillard E et al. 2016. Validation of Black Sea bass, Centropristis striata, Ages Using Oxytetracycline Marking and Scale Margin Increments. Appendix A1, Stock Assessment Report of Black Sea Bass. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 18p.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Stock Assessment Report of Black Sea Bass. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 247p.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Stock Assessment Summary Report of Black Sea Bass. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 8p.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Port-Based Black Sea Bass Outreach Project. Appendix A2, Stock Assessment Report of Black Sea Bass. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 21p.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Investigating the utility of inshore trawl surveys for developing black sea bass abundance indices. Appendix A3, Stock Assessment Report of Black Sea Bass. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 27p.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Fishery Management History. Appendix A4, Stock Assessment Report of Black Sea Bass. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 15p.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Term of Reference 1 – Spatial Issues. Appendix A5, Stock Assessment Report of Black Sea Bass. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 46p.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Black sea bass distribution maps – Distribution of State and Federal surveys and NEFSC spring survey distribution maps, 1989-2015. Appendix A7, Stock Assessment Report of Black Sea Bass. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 206p.

Working Group, Stock Assessment Workshop (SAW 62). 2016. ALK simulation to test efficacy of multinomial approach. Appendix A8, Stock Assessment Report of Black Sea Bass. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 19p.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Overall Model of All Plots. Appendix A10, Stock Assessment Report of Black Sea Bass. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 156p.

Working Group, Stock Assessment Workshop (SAW 62). 2016. North Model of All Plots. Appendix A11, Stock Assessment Report of Black Sea Bass. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 134p.

Working Group, Stock Assessment Workshop (SAW 62). 2016. South Model of All Plots. Appendix A12, Stock Assessment Report of Black Sea Bass. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 116p.

Working Group, Stock Assessment Workshop (SAW 62). 2016. North Area Exchange of All Plots. Appendix A13, Stock Assessment Report of Black Sea Bass. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 134p.

Working Group, Stock Assessment Workshop (SAW 62). 2016. South Area Exchange of All Plots. Appendix A14, Stock Assessment Report of Black Sea Bass. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 116p.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Two Area Model Justification. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 1p.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Combined ASAP Retros. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. Power Point presentation. 3 slides.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Comparison of results for Black Sea Bass ASAP Two Area model and SS (run 134). SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 1p.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Groundfish retro-adjusted values used in management. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 1p.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Index tables. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 2p.

Working Group, Stock Assessment Workshop (SAW 62). 2016. M Profile Obj FX Components. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 2p.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Model Justification Diagnostics. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. Power Point presentation. 8 slides.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Normalized indices used in both North and South area models. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. Power Point presentation. 4 slides.

Working Group, Stock Assessment Workshop (SAW 62). 2016. SS comparisons. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 1p.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Standardized Age Comp Residual Plots. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. Power Point presentation. 18 slides.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Stock recruit. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. Power Point presentation. 2 slides.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Black sea bass Z-score normalized index values. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. Power Point presentation. 1 slide.

Presentations

Working Group, Stock Assessment Workshop (SAW 62). 2016. Black Sea Bass Assessment Review. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. Power Point presentation. 261 slides.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Commercial VTRs. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. Power Point presentation. 40 slides.

Working Group, Stock Assessment Workshop (SAW 62). 2016. VTR Trawl and Spring Survey. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. Power Point presentation. 23 slides.

Witch Flounder

Background Papers

Butterworth DS and Rademeyer RA. 2016. Further Remarks on Gulf of Maine-Georges Bank Witch Flounder Assessment Results. Working Paper for SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 12p.

Cadrin S and Wright B. 2016. Fishery Catch Rates of Working Flounder. Working Paper for SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 17p.

DeCelles G. 2016. An Assessment of Witch Flounder (*Glyptocephalus cynoglossus*) Stock Structure. Working Paper for SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 18p.

Friedland K. 2016. Data to inform habitat model construction for witch flounder. Working Paper for SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 23p.

Friedland K. 2016. Estimated witch flounder habitat using random forest models. Working Paper for SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 15p.

Hare J et al. 2016. In situ temperature and salinity data for use in stock assessments. Working Paper for SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 4p.

Hare J et al. 2016. Empirical estimates of maximum catchability of Witch Flounder *Glyptocaphalus cynoglossus L.* on the Northeast Fisheries Science Center Fall bottom trawl survey. Working Paper for SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 27p.

Hare J et al. 2016. Environmentally explicit stock-recruitment relationships in Witch Flounder. Working Paper for SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 5p.

Kritzer JP et al. 2016. Spatial and Temporal Patterns in Habitat Use and Depth Distribution of Witch Flounder: Implications for Stock Assessment. Working Paper for SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 11p.

Northeast Fisheries Science Center. 2008. Assessment of 19 Northeast Groundfish Stocks through 2007: Report of the 3rd Groundfish Assessment Review Meeting (GARM III), Northeast Fisheries Science Center, Woods Hole, Massachusetts, August 4-8, 2008. US Dep Commer, NOAA Fisheries, Northeast Fish Sci Cent Ref Doc. 08-15; 884 p + xvii.

Northeast Fisheries Science Center. 2012. Assessment or Data Updates of 13 Northeast Groundfish Stocks through 2010. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 12-06; 789 p.

Northeast Fisheries Science Center. 2015. Operational Assessment of 20 Northeast Groundfish Stocks, Updated Through 2014. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 15-24; 251 p.

Odell J et al. 2016. NSC-AFM-GFCPF Witch Flounder Letter. Working Paper for SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 2p.

Palmer MC. 2016. Catch curve analysis of witch flounder fishery and survey catch-at-age data. Working Paper for SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 9p.

Richardson D. 2016. A minimum estimate of Witch Flounder spawning stock biomass using experimental estimates of catchability on the NEFSC trawl survey. Working Paper for SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 4p.

Terceiro M. 2016. TOR 1: Description of commercial fishery Dealer Report trawl gear landings and effort and modeling landings rate (LPUE) data for witch flounder. Working Paper for SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 21p.

Terceiro M. 2016. TOR 1: Description of commercial fishery Dealer Report trawl gear landings and effort and modeling landings rate (LPUE) data for witch flounder: 'Directed' Trips (=>40% of trip landings). Working Paper for SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 11p.

Terceiro M. 2016. TOR 1 & 2: Modeling commercial fishery Dealer Report fish trawl gear landings rate (LPUE) data for witch flounder: 'Directed' Trips (=>40%, =>25%, and =>10% of trip landings). Working Paper for SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 27p.

Terceiro M. 2016. TOR 1: Description and modeling of NEFOP (Observer) fish trawl gear catch rate (CPUE) data for witch flounder. Working Paper for SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 17p.

Terceiro M. 2016. TOR 1 & 2: Description of Vessel Trip Report trawl gear catch and effort data and modeling catch rates (CPUE) for witch flounder. Working Paper for SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 30p.

Walsh HJ et al. 2016. Changes in the distributions of larval, juvenile, and adult witch flounder in the Northeast US Shelf Ecosystem: Updates Through 2015. Working Paper for SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 9p.

Wigley SE. 2016. Rough vs Smooth Bottom Type: An Initial Exploration. Working Paper for SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 15p.

Wigley SE and Burnett JM. 2016. Preliminary Estimates of Biological and Yield Characteristics of Deep-water Witch Flounder (Glyptocephalus cynoglossus) in the Georges Bank-Mid-Atlantic Bight Region. J Northw Atl Fish Sci. 31:181-194.

Wigley SE. 2016. Refinements to 1982-2014 Witch Flounder Discard Estimates. Working Paper for SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 39p.

Wigley SE, Brodziak JKT,Col L. 2003. Assessment of the Gulf of Maine and Georges Bank witch flounder stock for 2003. Northeast Fish. Sci. Cent. Ref. Doc. 03-14; 186 p.

Working Papers

Butterworth DS and Rademeyer RA. 2016. Response to reviewer requests in regard to the impact of selectivity doming in the preferred SCAA model. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 4p.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Stock Assessment of Witch Flounder for 2016. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 523p.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Stock Assessment Summary of Witch Flounder for 2016. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 16p.

Presentations

Hare J et al. 2016. Empirical Estimates of Maximum Catchability of Witch Flounder on the Northeast Fisheries Science Center Fall Bottom Trawl Survey. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. Power Point presentation. 23 slides.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Witch Flounder Assessment Review, TORs 1-3. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. Power Point presentation. 89 slides.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Witch Flounder Assessment Review, TORs 4-9. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. Power Point presentation. 123 slides.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Witch Flounder SARC Discussion Slides. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. Power Point presentation. 32 slides.

Appendix 3: Statement of Work

Statement of Work

National Oceanic and Atmospheric Administration (NOAA)
National Marine Fisheries Service (NMFS)
Center for Independent Experts (CIE) Program
External Independent Peer Review

62nd Stock Assessment Workshop/Stock Assessment Review Committee (SAW/SARC)
Benchmark stock assessment for Black sea bass and Witch flounder

Background

The National Marine Fisheries Service (NMFS) is mandated by the Magnuson-Stevens Fishery Conservation and Management Act, Endangered Species Act, and Marine Mammal Protection Act to conserve, protect, and manage our nation's marine living resources based upon the best scientific information available (BSIA). NMFS science products, including scientific advice, are often controversial and may require timely scientific peer reviews that are strictly independent of all outside influences. A formal external process for independent expert reviews of the agency's scientific products and programs ensures their credibility. Therefore, external scientific peer reviews have been and continue to be essential to strengthening scientific quality assurance for fishery conservation and management actions.

Scientific peer review is defined as the organized review process where one or more qualified experts review scientific information to ensure quality and credibility. These expert(s) must conduct their peer review impartially, objectively, and without conflicts of interest. Each reviewer must also be independent from the development of the science, without influence from any position that the agency or constituent groups may have. Furthermore, the Office of Management and Budget (OMB), authorized by the Information Quality Act, requires all federal agencies to conduct peer reviews of highly influential and controversial science before dissemination, and that peer reviewers must be deemed qualified based on the OMB Peer Review Bulletin standards.
(http://www.cio.noaa.gov/services_programs/pdfs/OMB_Peer_Review_Bulletin_m05-03.pdf).
Further information may be obtained from www.ciereviews.org.

Scope

The Northeast Regional Stock Assessment Review Committee (SARC) meeting is a formal, multiple-day meeting of stock assessment experts who serve as a panel to peer-review tabled stock assessments and models. The SARC peer review is the cornerstone of the Northeast Stock Assessment Workshop (SAW) process, which includes assessment development and report preparation (which is done by SAW Working Groups or ASMFC technical committees), assessment peer review (by the SARC), public presentations, and document publication. This review determines whether or not the scientific assessments are adequate to serve as a basis for developing fishery management advice. Results provide the scientific basis for fisheries within the jurisdiction of NOAA's Greater Atlantic Regional Fisheries Office (GARFO).

The purpose of this meeting will be to provide an external peer review of a benchmark stock assessment for **Black sea bass and Witch flounder**. The requirements for the peer review follow. This Statement of Work (SOW) also includes Appendix 1: TORs for the stock assessment, which are the responsibility of the analysts; Appendix 2: a draft meeting agenda; Appendix 3: Individual Independent Review Report Requirements; and Appendix 4: SARC Summary Report Requirements.

Requirements

NMFS requires three reviewers under this contract (i.e. subject to CIE standards for reviewers) to participate in the panel review. The SARC chair, who is in addition to the three reviewers, will be provided by either the New England or Mid-Atlantic Fishery Management Council's Science and Statistical Committee; although the SARC chair will be participating in this review, the chair's participation (i.e. labor and travel) is not covered by this contract.

Each reviewer will write an individual review report in accordance with the SOW, OMB Guidelines, and the TORs below. All TORs must be addressed in each reviewer's report. No more than one of the reviewers selected for this review is permitted to have served on a SARC panel that reviewed this same species in the past. The reviewers shall have working knowledge and recent experience in the application of modern fishery stock assessment models. Expertise should include forward projecting statistical catch-at-age models. Reviewers should also have experience in evaluating measures of model fit, identification, uncertainty, and forecasting. Reviewers should have experience in development of Biological Reference Points (BRPs) that includes an appreciation for the varying quality and quantity of data available to support estimation of BRPs. For Black sea bass, knowledge of spatial models and complex fisheries with multiple fleets and recreational fisheries would be useful. For Witch flounder, knowledge of flatfish ecology would be useful.

Requirements for Reviewers

- Review the background materials and reports prior to the review meeting
- Attend and participate in the panel review meeting
    - The meeting will consist of presentations by NOAA and other scientists, stock assessment authors and others to facilitate the review, to provide any additional information required by the reviewers, and to answer any questions from reviewers
- Reviewers shall conduct an independent peer review in accordance with the requirements specified in this SOW and TORs, in adherence with the required formatting and content guidelines; reviewers are not required to reach a consensus.
- Each reviewer shall assist the SARC Chair with contributions to the SARC Summary Report
- Deliver individual Independent Review Reports to the Government according to the specified milestone dates

- This report should explain whether each stock assessment Term of Reference of the SAW was or was not completed successfully during the SARC meeting, using the criteria specified below in the "Requirements for SARC panel."
- If any existing Biological Reference Points (BRP) or their proxies are considered inappropriate, the Independent Report should include recommendations and justification for suitable alternatives. If such alternatives cannot be identified, then the report should indicate that the existing BRPs are the best available at this time.
- During the meeting, additional questions that were not in the Terms of Reference but that are directly related to the assessments may be raised. Comments on these questions should be included in a separate section at the end of the Independent Report produced by each reviewer.
- The Independent Report can also be used to provide greater detail than the SARC Summary Report on specific stock assessment Terms of Reference or on additional questions raised during the meeting.

Requirements for SARC panel

- During the SARC meeting, the panel is to determine whether each stock assessment Term of Reference (TOR) of the SAW was or was not completed successfully. To make this determination, panelists should consider whether the work provides a scientifically credible basis for developing fishery management advice. Criteria to consider include: whether the data were adequate and used properly, the analyses and models were carried out correctly, and the conclusions are correct/reasonable. If alternative assessment models and model assumptions are presented, evaluate their strengths and weaknesses and then recommend which, if any, scientific approach should be adopted. Where possible, the SARC chair shall identify or facilitate agreement among the reviewers for each stock assessment TOR of the SAW.
- If the panel rejects any of the current BRP or BRP proxies (for B_{MSY} and F_{MSY} and MSY), the panel should explain why those particular BRPs or proxies are not suitable, and the panel should recommend suitable alternatives. If such alternatives cannot be identified, then the panel should indicate that the existing BRPs or BRP proxies are the best available at this time.
- Each reviewer shall complete the tasks in accordance with the SOW and Schedule of Milestones and Deliverables below.

Requirements for SARC chair and reviewers combined:

Review both the Assessment Report and the draft Assessment Summary Report. The draft Assessment Summary Report is reviewed and edited to assure that it is consistent with the outcome of the peer review, particularly statements that address stock status and assessment uncertainty.

The SARC Chair, with the assistance from the reviewers, will write the SARC Summary Report. Each reviewer and the chair will discuss whether they hold similar views on each stock assessment Term of Reference and whether their opinions can be summarized into a single conclusion for all or only for some of the Terms of Reference of the SAW. For terms where a similar view can be reached, the SARC Summary Report will contain a summary of such opinions. In cases where multiple and/or differing views exist on a given Term of Reference, the SARC Summary Report will note that there is no agreement and will specify - in a summary manner – what the different opinions are and the reason(s) for the difference in opinions.

The chair's objective during this SARC Summary Report development process will be to identify or facilitate the finding of an agreement rather than forcing the panel to reach an agreement. The chair will take the lead in editing and completing this report. The chair may express the chair's opinion on each Term of Reference of the SAW, either as part of the group opinion, or as a separate minority opinion. The SARC Summary Report will not be submitted, reviewed, or approved by the Contractor.

If any existing Biological Reference Points (BRP) or BRP proxies are considered inappropriate, the SARC Summary Report should include recommendations and justification for suitable alternatives. If such alternatives cannot be identified, then the report should indicate that the existing BRP proxies are the best available at this time.

Foreign National Security Clearance

When reviewers participate during a panel review meeting at a government facility, the NMFS Project Contact is responsible for obtaining the Foreign National Security Clearance approval for reviewers who are non-US citizens. For this reason, the reviewers shall provide requested information (e.g., first and last name, contact information, gender, birth date, country of birth, country of citizenship, country of permanent residence, country of current residence, dual citizenship (yes, no), passport number, country of passport, travel dates.) to the NEFSC SAW Chair for the purpose of their security clearance, and this information shall be submitted at least 30 days before the peer review in accordance with the NOAA Deemed Export Technology Control Program NAO 207-12 regulations available at the Deemed Exports NAO website: http://deemedexports.noaa.gov/ and http://deemedexports.noaa.gov/compliance_access_control_procedures/noaa-foreign-national-registration-system.html. The contractor is required to use all appropriate methods to safeguard Personally Identifiable Information (PII).

Place of Performance

The place of performance shall be at the contractor's facilities, and at the Northeast Fisheries Science Center in Woods Hole, Massachusetts.

Period of Performance

The period of performance shall be from the time of award through January 15, 2017. Each reviewer's duties shall not exceed 14 days to complete all required tasks.

Schedule of Milestones and Deliverables: The contractor shall complete the tasks and deliverables in accordance with the following schedule.

No later than November 15, 2016	Contractor sends reviewer contact information to the COR, who then sends this to the NMFS Project Contact
No later than November 15, 2016	NMFS Project Contact will provide reviewers the pre-review documents
Nov. 29 – Dec. 2, 2016	Each reviewer participates and conducts an independent peer review during the panel review meeting in Woods Hole, MA
December 2, 2016	SARC Chair and reviewers work at drafting reports during meeting at Woods Hole, MA, USA
December 16, 2016	Reviewers submit draft independent peer review reports to the contractor's technical team for review
December 16, 2016	Draft of SARC Summary Report, reviewed by all reviewers, due to the SARC Chair *
December 23, 2016	SARC Chair sends Final SARC Summary Report, approved by reviewers, to NMFS Project contact (i.e., SAW Chairman)
December 30, 2016	Contractor submits independent peer review reports to the COR and technical point of contact (POC)
January 6, 2017	The COR and/or technical POC distributes the final reports to the NMFS Project Contact and regional Center Director

* The SARC Summary Report will not be submitted to, reviewed, or approved by the Contractor.

Applicable Performance Standards

The acceptance of the contract deliverables shall be based on three performance standards: (1) The reports shall be completed in accordance with the required formatting and content; (2) The reports shall address each TOR as specified; (3) The reports shall be delivered as specified in the schedule of milestones and deliverables.

Travel

All travel expenses shall be reimbursable in accordance with Federal Travel Regulations (http://www.gsa.gov/portal/content/104790). International travel is authorized for this contract. Travel is not to exceed $20,000.

Restricted or Limited Use of Data

The contractors may be required to sign and adhere to a non-disclosure agreement.

Project Contacts

Dr. James Weinberg, NEFSC SAW Chair Northeast
Fisheries Science Center
166 Water Street, Woods Hole, MA 02543
James.Weinberg@noaa.gov Phone: 508-495-2352

Dr. William Karp, NEFSC Science Director
Northeast Fisheries Science Center 166 Water St.,
Woods Hole, MA 02543
william.karp@noaa.gov Phone: 508-495-2233

Appendix 1. Terms of Reference for the SAW Working Group (62nd SAW/SARC Stock Assessment)

T*he SARC Review Panel shall assess whether or not the SAW Working Group has reasonably and satisfactorily completed the following actions.*

A. Black sea bass

1. Summarize the conclusions of the February 2016 SSC peer review regarding the potential for spatial partitioning of the black sea bass stock. The consequences for the stock assessment will be addressed in TOR-6.)

2. Estimate catch from all sources including landings and discards. Characterize the uncertainty in these sources of data. Evaluate available information on discard mortality and, if appropriate, update mortality rates applied to discard components of the catch. Describe the spatial and temporal distribution of fishing effort.

3. Present the survey data being used in the assessment (e.g., indices of abundance, recruitment, state surveys, age-length data, etc.). Investigate the utility of fishery dependent indices as a measure of relative abundance. Characterize the uncertainty and any bias in these sources of data.

4. Consider the consequences of environmental factors on the estimates of abundance or relative indices derived from surveys.

5. Investigate implications of hermaphroditic life history on stock assessment model. If possible, incorporate parameters to account for hermaphroditism.

6. Estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock), using measures that are appropriate to the assessment model, for the time series (integrating results from TORs-1,-4, & -5 as appropriate), and estimate their uncertainty. Include a historical retrospective analysis and past projection performance evaluation to allow a comparison with most recent assessment results.

7. Estimate biological reference points (BRPs; point estimates or proxies for B_{MSY}, $B_{THRESHOLD}$, F_{MSY}, and MSY), including defining BRPs for spatially explicit areas if appropriate, and provide estimates of their uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for BRPs. Comment on the appropriateness of existing BRPs and the “new” (i.e., updated, redefined, or alternative) BRPs.

8. Evaluate overall stock status with respect to a new model or new models that considered spatial units developed for this peer review.

9. Develop approaches and apply them to conduct stock projections.
    a. Provide numerical annual projections (3-5 years) and the statistical distribution (e.g., probability density function) of the OFL (overfishing level) that fully incorporates observation, process and model uncertainty (see Appendix to the SAW TORs). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. Use a sensitivity analysis approach in which a range of assumptions about the most important uncertainties in the assessment are considered (e.g., terminal year abundance, variability in recruitment, and definition of BRPs for black sea bass).
    b. Comment on which projections seem most realistic. Consider major uncertainties in the assessment as well as the sensitivity of the projections to various assumptions.
    c. Describe this stock's vulnerability (see "Appendix to the SAW TORs") to becoming overfished, and how this could affect the choice of ABC.

10. Review, evaluate and report on the status of the SARC and Working Group research recommendations listed in recent SARC reviewed assessments and review panel reports. Identify new research recommendations.

B. Witch flounder

1. Estimate catch from all sources including landings and discards. Describe the spatial and temporal distribution of landings, discards, and fishing effort. Characterize the uncertainty in these sources of data.

2. Present available federal, state, and other survey data, indices of relative or absolute abundance, recruitment, etc. Characterize the uncertainty and any bias in these sources of data and compare survey coverage to locations of fishery catches. Select the surveys and indices for use in the assessment.

3. Investigate effects of environmental factors and climate change on recruitment, growth and natural mortality of witch flounder. If quantifiable relationships are identified, consider incorporating these into the stock assessment.

4. Estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series (integrating results from TOR-3 if appropriate), and estimate their uncertainty. Include a historical retrospective analysis to allow a comparison with previous assessment results and previous projections. Compare F's and SSB's that were projected during the previous assessment to their realized values.

5. State the existing stock status definitions for "overfished" and "overfishing". Then update or redefine biological reference points (BRPs; point estimates or proxies for B_{MSY}, $B_{THRESHOLD}$, F_{MSY} and MSY) and provide estimates of their uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for BRPs. Comment on the scientific adequacy of existing BRPs and the "new" (i.e., updated, redefined, or alternative) BRPs.

6. Evaluate stock status with respect to the existing model (from previous peer reviewed accepted assessment) and with respect to a new model (or possibly models, in accord with guidance in attached "Appendix to the SAW Assessment TORs") developed for this peer review. In both cases, evaluate whether the stock is rebuilt.
    a. When working with the existing model, update it with new data and evaluate stock status (overfished and overfishing) with respect to the updated BRP estimates.
    b. Then use the newly proposed model (or possibly models, in accord with guidance in "Appendix to the SAW Assessment TORs") and evaluate stock status with respect to "new" BRPs and their estimates (from TOR-5).

7. Develop approaches and apply them to conduct stock projections.
    a. Provide numerical annual projections (3 years) and the statistical distribution (e.g., probability density function) of the catch at F_{MSY} or an F_{MSY} proxy (i.e. the overfishing level, OFL) (see Appendix). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. Use a sensitivity analysis approach in which a range of assumptions about the most important uncertainties in the assessment are considered (e.g., terminal year abundance, magnitude and variability in recruitment).
    b. Comment on which projections seem most realistic. Consider the major uncertainties in the assessment as well as sensitivity of the projections to various assumptions. Identify reasonable projection parameters (recruitment, weight-at-age, retrospective adjustments, etc.) to use when setting specifications.
    c. Describe this stock's vulnerability to becoming overfished, and how this could affect the choice of ABC. The choice takes scientific uncertainty into account (see Appendix).

8. Evaluate the validity of the current stock definition, taking into account what is known about migration, and make a recommendation about whether there is a need to modify the current stock definition for future stock assessments.

9. Review, evaluate and report on the status of research recommendations from the last peer reviewed benchmark stock assessment. Identify new research recommendations.

Clarification of Terms

used in the SAW/SARC Terms of Reference

Guidance to SAW WG about "Number of Models to include in the Assessment Report":
In general, for any TOR in which one or more models are explored by the WG, give a detailed presentation of the "best" model, including inputs, outputs, diagnostics of model adequacy, and sensitivity analyses that evaluate robustness of model results to the assumptions. In less detail, describe other models that were evaluated by the WG and explain their strengths, weaknesses and results in relation to the "best" model. If selection of a "best" model is not possible, present alternative models in detail, and summarize the relative utility each model, including a comparison of results. It should be highlighted whether any models represent a minority opinion.

On "Acceptable Biological Catch" (DOC Nat. Stand. Guidel. Fed. Reg., v. 74, no. 11, 1-16-2009):

Acceptable biological catch (ABC) is a level of a stock or stock complex's annual catch that accounts for the scientific uncertainty in the estimate of Overfishing Limit (OFL) and any other scientific uncertainty…" *(p. 3208) [In other words, OFL ≥ ABC.]*

ABC for overfished stocks. For overfished stocks and stock complexes, a rebuilding ABC must be set to reflect the annual catch that is consistent with the schedule of fishing mortality rates in the rebuilding plan. *(p. 3209)*

NMFS expects that in most cases ABC will be reduced from OFL to reduce the probability that overfishing might occur in a year. (p. 3180)

ABC refers to a level of ''catch'' that is ''acceptable'' given the ''biological'' characteristics of the stock or stock complex. As such, Optimal Yield (OY) does not equate with ABC. The specification of OY is required to consider a variety of factors, including social and economic factors, and the protection of marine ecosystems, which are not part of the ABC concept. (p. 3189)

On "Vulnerability" (DOC Natl. Stand. Guidelines. Fed. Reg., v. 74, no. 11, 1-16-2009):

"Vulnerability. A stock's vulnerability is a combination of its productivity, which depends upon its life history characteristics, and its susceptibility to the fishery. Productivity refers to the capacity of the stock to produce Maximum Sustainable Yield (MSY) and to recover if the population is depleted, and susceptibility is the potential for the stock to be impacted by the fishery, which includes direct captures, as well as indirect impacts to the fishery (e.g., loss of habitat quality)." (p. 3205)

Participation among members of a Stock Assessment Working Group:

Anyone participating in SAW meetings that will be running or presenting results from an assessment model is expected to supply the source code, a compiled executable, an input file with the proposed configuration, and a detailed model description in advance of the model meeting. Source code for NOAA Toolbox programs is available on request. These measures allow transparency and a fair evaluation of differences that emerge between models.

Appendix 3. Individual Independent Peer Review Report Requirements

1. The independent peer review report shall be prefaced with an Executive Summary providing a concise summary of whether they accept or reject the work that they reviewed, with an explanation of their decision (strengths, weaknesses of the analyses, etc.).

2. The report must contain a background section, description of the individual reviewers' roles in the review activities, summary of findings for each TOR in which the weaknesses and strengths are described, and conclusions and recommendations in accordance with the TORs. The independent report shall be an independent peer review, and shall not simply repeat the contents of the SARC Summary Report.

   a. Reviewers should describe in their own words the review activities completed during the panel review meeting, including a concise summary of whether they accept or reject the work that they reviewed, and explain their decisions (strengths, weaknesses of the analyses, etc.), conclusions, and recommendations.

   b. Reviewers should discuss their independent views on each TOR even if these were consistent with those of other panelists, but especially where there were divergent views.

   c. Reviewers should elaborate on any points raised in the SARC Summary Report that they believe might require further clarification.

   d. The report may include recommendations on how to improve future assessments.

3. The report shall include the following appendices:

Appendix 1: Bibliography of materials provided for review
Appendix 2: A copy of this Statement of Work
Appendix 3: Panel membership or other pertinent information from the panel review meeting.

Appendix 4. SARC Summary Report Requirements

1. The main body of the report shall consist of an introduction prepared by the SARC chair that will include the background and a review of activities and comments on the appropriateness of the process in reaching the goals of the SARC. Following the introduction, for each assessment reviewed, the report should address whether or not each Term of Reference of the SAW Working Group was completed successfully. For each Term of Reference, the SARC Summary Report should state why that Term of Reference was or was not completed successfully.

   To make this determination, the SARC chair and reviewers should consider whether or not the work provides a scientifically credible basis for developing fishery management advice. If the reviewers and SARC chair do not reach an agreement on a Term of Reference, the report should explain why. It is permissible to express majority as well as minority opinions.

   The report may include recommendations on how to improve future assessments.

2. If any existing Biological Reference Points (BRPs) or BRP proxies are considered inappropriate, include recommendations and justification for alternatives. If such alternatives cannot be identified, then indicate that the existing BRPs or BRP proxies are the best available at this time.

3. The report shall also include the bibliography of all materials provided during the SAW, and relevant papers cited in the SARC Summary Report, along with a copy of the Statement of Work.

   The report shall also include as a separate appendix the assessment Terms of Reference used for the SAW, including any changes to the Terms of Reference or specific topics/issues directly related to the assessments and requiring Panel advice.